

WE CLAIM:

1. A method, comprising:

estimating the noise power spectrum for each frame of an audio signal based on a plurality of signal power spectrum values computed from a corresponding plurality of adjacent frames;

computing dynamically an over-subtraction factor for each frame of the audio signal based on the estimated noise power spectrum of the frame;

reducing the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor computed for the frame.

2. The method according to claim 1, wherein said estimating the noise power spectrum comprises:

computing the signal energy for each sub frequency band of each frame of the audio signal;

deriving noise energy for each subband of each frame based on a plurality of signal energy values computed with respect to the same subband for a plurality of corresponding frames.

3. The method according to claim 2, wherein deriving the noise energy includes:

taking a minimum signal energy of each subband across a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame;

computing an average signal energy of a set of pre-determined percentage of the smallest signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame; and

taking a signal energy value corresponding to a pre-determined percentile of the signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame.

4. The method according to claim 1, wherein said computing the over-subtraction factor comprises:

determining the signal to noise ratio of each frame based on the corresponding signal power spectrum and noise power spectrum computed and estimated for the frame; and

deriving an over-subtraction factor for the frame based on the signal to noise ratio dynamically determined for the frame.

5. The method according to claim 4, wherein:

the signal to noise ratio of the frame is computed as

$$SNR(r) = 10 \log \left(\frac{\sum_w P_y(r, w) - \sum_w P_n(r, w)}{\sum_w P_n(r, w)} \right)$$

where $SNR(r)$ represents the signal to noise ratio estimated for frame r , $P_y(r, w)$ represents signal energy of frame r at subband w , and $P_n(r, w)$ represents noise energy of frame r at subband w ; and

the over-subtraction factor for the frame is computed based on the signal to noise ratio as:

$$OSF(r) = \frac{\varepsilon}{1 + \eta SNR(r)}$$

where $OSF(r)$ represents the over-subtraction factor for frame r and ε and η are pre-determined parameters.

6. The method according to claim 5, wherein said subtracting comprises:

computing a subtraction amount for each subband of each frame using the corresponding over-subtraction factor computed for the frame, the signal energy computed for the subband of the frame, and the noise energy computed for the subband of the frame; and

subtracting the signal energy of the subband of the frame by the subtraction amount according to the following rule:

$$P_s(r, w) = \begin{cases} P_y(r, w) - OSF(r) \times P_n(r, w) & \text{if } P_y(r, w) - OSF(r) \times P_n(r, w) > 0 \\ \sigma & \text{if } P_y(r, w) - OSF(r) \times P_n(r, w) \leq 0 \end{cases}$$

where $P_s(r, w)$ represents the subtracted signal energy at subband w of frame r and σ is a pre-determined constant.

7. The method according to claim 1, further comprising:

performing a Fourier transform on the audio signal prior to said estimating the noise power spectrum to produce a transformed signal based on which the signal power spectrum of the audio signal is computed; and

performing a corresponding inverse Fourier transform, after said subtracting, using the subtracted signal power spectrum to produce an enhanced audio signal.

8. A method, comprising:

receiving an audio signal;

enhancing the audio signal to produce an enhanced audio signal via spectral subtraction using an over-subtraction amount dynamically computed based on the noise power spectrum of the audio signal estimated for each frame of the audio signal based on a plurality of signal power spectrum values of the audio signal computed from a corresponding plurality of adjacent frames; and

utilizing the enhanced audio signal.

9. The method according to claim 8, wherein said enhancing comprises:

performing a Fourier transform on the received audio signal to produce a transformed signal;

estimating, based on the transformed signal, noise power spectrum for each frame of the audio signal based on a plurality of signal power spectrum values computed from a corresponding plurality of adjacent frames of the audio signal;

computing dynamically an over-subtraction factor for each frame of the audio signal based on signal to noise ratio computed for the frame based on the signal power spectrum and the noise power spectrum of the frame;

performing spectral subtraction of the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor computed for the frame to produce subtracted signal power spectrum; and

performing an inverse Fourier transform based on the subtracted signal power spectrum to produce the enhanced audio signal.

10. The method according to claim 9, wherein said estimating the noise power spectrum includes:

taking a minimum signal energy of each subband across a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame;

computing an average signal energy of a set of pre-determined percentage of the smallest signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame; and

taking a signal energy value corresponding to a pre-determined percentile of the signal energy values of the subband from a pre-determined

plurality of adjacent frames as the estimated noise energy of the subband for the frame.

11. The method according to claim 8, wherein said utilizing includes:

playing back the enhanced audio signal;

performing speaker identification based on the enhanced audio signal;

segmenting the audio signal based on the enhanced audio signal; and

performing speech recognition on the enhanced audio signal.

12. The method according to claim 8, wherein said enhancing is an embedded operation of said utilizing.

13. A system, comprising:

a dynamic noise power spectrum estimation mechanism configured to estimate noise power spectrum using at least one signal power spectrum value of the audio signal computed for a corresponding plurality of adjacent frames of the audio signal;

an over-subtraction factor estimation mechanism configured to dynamically compute an over-subtraction factor for each frame of the audio signal based on the noise power spectrum estimated for the frame; and

a spectral subtraction mechanism configured to reduce the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor dynamically computed for the frame.

14. The system according to claim 13, wherein the dynamic noise power spectrum estimation mechanism comprises:

a signal power spectrum estimator configured to compute the signal energy for each sub frequency band of each frame; and

a noise power spectrum estimator configured to derive noise energy for each subband of each frame based on a plurality of signal energies at the same subband computed for a corresponding plurality of adjacent frames, wherein the noise energy is computed as one of a minimum signal energy at each subband across a pre-determined number of adjacent frames.

15. The system according to claim 14, wherein the noise energy is computed as one of an average signal energy, averaged over a set of pre-determined smallest signal energy values at the subband computed from a pre-determined number of adjacent frames, and a signal energy corresponding to a pre-determined percentile across a pre-determined number of adjacent frames.

16. The system according to claim 13, wherein the over-subtraction factor estimation mechanism comprises:

a dynamic signal to noise ratio estimator configured to determine a signal to noise ratio for each frame based on the corresponding signal power spectrum and noise power spectrum computed and estimated for the frame; and

an over-subtraction factor estimator configured to derive an over-subtraction factor for each frame based on the signal to noise ratio determined for the frame.

17. The system according to claim 13, further comprising:

a preprocessing mechanism configured to perform a Fourier transform on the audio signal to produce a transformed signal based on which the signal power spectrum is computed; and

an inverse Fourier transform mechanism configured to performing an inverse Fourier transform using the subtracted signal power spectrum to produce an enhanced audio signal.

18. A system, comprising:

a spectral subtraction based audio enhancer configured to enhance an audio signal to produce an enhanced audio signal via spectral subtraction using a subtraction amount dynamically computed based on noise power spectrum of the audio signal dynamically estimated based on at least one signal power spectrum value of the audio signal computed from a corresponding plurality of adjacent frames; and

an audio signal processing mechanism configured to utilizing the enhanced audio signal.

19. The system according to claim 18, wherein the spectral subtraction based audio enhancer comprises:

a preprocessing mechanism configured to perform a Fourier transform on the audio signal to produce a transformed signal;

a dynamic noise power spectrum estimation mechanism configured to estimate, based on the transformed signal, noise power spectrum using at least one signal power spectrum values of the audio signal computed for a corresponding plurality of adjacent frames of the audio signal;

an over-subtraction factor estimation mechanism configured to dynamically compute an over-subtraction factor for each frame of the audio signal based on dynamic signal to noise ratio of the frame estimated based on the noise power spectrum estimated for the frame; and

a spectral subtraction mechanism configured to reduce the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor dynamically determined for the frame; and

an inverse Fourier transform mechanism configured to performing an inverse Fourier transform using the subtracted signal power spectrum to produce an enhanced audio signal.

20. The system according to claim 18, wherein the spectral subtraction based audio enhancer is embedded in the audio signal processing mechanism.

21. An article comprising a storage medium having stored thereon instructions that, when executed by a machine, result in the following:

estimating the noise power spectrum for each frame of an audio signal based on a plurality of signal power spectrum values computed from a corresponding plurality of adjacent frames;

computing dynamically an over-subtraction factor for each frame of the audio signal based on the estimated noise power spectrum of the frame;

reducing the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor computed for the frame.

22. The article according to claim 21, wherein said estimating the noise power spectrum comprises:

computing the signal energy for each sub frequency band of each frame of the audio signal;

deriving noise energy for each subband of each frame based on a plurality of signal energy values computed with respect to the same subband for a plurality of corresponding frames.

23. The article according to claim 22, wherein said deriving the noise energy includes:

taking a minimum signal energy of each subband across a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame;

computing an average signal energy of a set of pre-determined percentage of the smallest signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame; and

taking a signal energy value corresponding to a pre-determined percentile of the signal energy values of the subband from a pre-determined

plurality of adjacent frames as the estimated noise energy of the subband for the frame.

24. The article according to claim 21, wherein said computing the over-subtraction factor comprises:

determining the signal to noise ratio of each frame based on the corresponding signal power spectrum and noise power spectrum computed and estimated for the frame; and

deriving an over-subtraction factor for the frame based on the signal to noise ratio dynamically determined for the frame.

25. The article according to claim 24, wherein:

the signal to noise ratio of the frame is computed as

$$SNR(r) = 10 \log \left(\frac{\sum_w P_y(r, w) - \sum_w P_n(r, w)}{\sum_w P_n(r, w)} \right)$$

where $SNR(r)$ represents the signal to noise ratio estimated for frame r , $P_y(r, w)$ represents signal energy of frame r at subband w , and $P_n(r, w)$ represents noise energy of frame r at subband w ; and

the over-subtraction factor for the frame is computed based on the signal to noise ratio as:

$$OSF(r) = \frac{\varepsilon}{1 + \eta SNR(r)}$$

where $OSF(r)$ represents the over-subtraction factor for frame r and ϵ and η are pre-determined parameters.

26. The article according to claim 25, wherein said subtracting comprises:

computing a subtraction amount for each subband of each frame using the corresponding over-subtraction factor computed for the frame, the signal energy computed for the subband of the frame, and the noise energy computed for the subband of the frame; and

subtracting the signal energy of the subband of the frame by the subtraction amount according to the following rule:

$$P_s(r, w) = \begin{cases} P_y'(r, w) - OSF(r) \times P_n(r, w) & \text{if } P_y'(r, w) - OSF(r) \times P_n(r, w) > 0 \\ \sigma & \text{if } P_y'(r, w) - OSF(r) \times P_n(r, w) \leq 0 \end{cases}$$

where $P_s(r, w)$ represents the subtracted signal energy at subband w of frame r and σ is a pre-determined constant.

27. An article comprising a storage medium having stored thereon instructions that, when executed by a machine, result in the following:

receiving an audio signal;

enhancing the audio signal to produce an enhanced audio signal via spectral subtraction using an over-subtraction amount dynamically computed based on the noise power spectrum of the audio signal estimated for each frame of the audio signal based on a plurality of signal power spectrum values

of the audio signal computed from a corresponding plurality of adjacent frames; and

utilizing the enhanced audio signal.

28. The article according to claim 27, wherein said enhancing comprises:

performing a Fourier transform on the received audio signal to produce a transformed signal;

estimating, based on the transformed signal, noise power spectrum for each frame of the audio signal based on a plurality of signal power spectrum values computed from a corresponding plurality of adjacent frames of the audio signal;

computing dynamically an over-subtraction factor for each frame of the audio signal based on signal to noise ratio computed for the frame based on the signal power spectrum and the noise power spectrum of the frame;

performing spectral subtraction of the signal power spectrum of the audio signal at each frame in accordance with the over-subtraction factor computed for the frame to produce subtracted signal power spectrum; and

performing an inverse Fourier transform based on the subtracted signal power spectrum to produce the enhanced audio signal.

29. The article according to claim 28, wherein said estimating the noise power spectrum includes:

taking a minimum signal energy of each subband across a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame;

computing an average signal energy of a set of pre-determined percentage of the smallest signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame; and

taking a signal energy value corresponding to a pre-determined percentile of the signal energy values of the subband from a pre-determined plurality of adjacent frames as the estimated noise energy of the subband for the frame.